

Distributed Systems and Distributed Computing

Portions of this PPT draw from PPTs
by Fazli Amin

Outline

- Definition of Distributed Systems
- Characteristics of Distributed Systems
- Distributed Computing
- Hardware Concepts

2

Section 1

- Definition of Distributed Systems
- Characteristics of Distributed Systems
- Distributed Computing

3

Distributed Systems

- Definition:
 - The Collection of individual computers that appears to its users as a single coherent system.
- Coherent System:
 - The system in which all the individual systems shares the central memory (RAM).

4

Computer Clusters for Scalable Parallel Computing

- A computer cluster is a collection of interconnected stand-alone computers which can work together collectively and cooperatively as a single integrated computing resource pool.
- Clustering explores massive parallelism at the job level and achieves high availability (HA) through stand-alone operations.

5

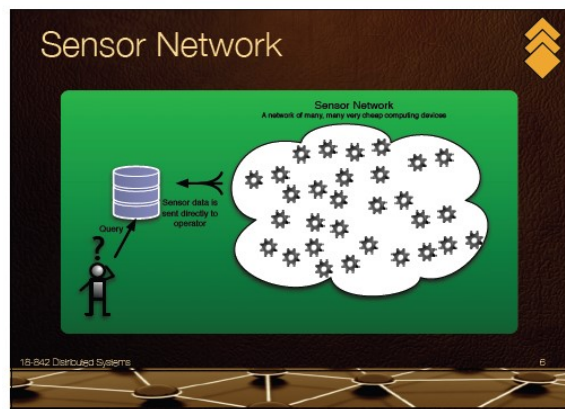
Examples

- Internet
- Aero plane Ticket Reservation System
- ATMs
- Mobile Cellular Phone Systems
- Centralized Data Base System Working in Banks
- LAN
- MAN
- WAN

6

DS: I know it when I see it

- -Big: Enterprise level
 - Examples of DS are all around us
- -Small: Personal
- -Innovative: Sensor Networks
- -Common: Web service



A Cluster @ CS Department, CSUF



9

A Cluster @ CS Department, CSUF



10

Biggest Challenges of DS

- • **Configuration:** How to find each other, get started, etc
- • **Consensus:** How to agree on stuff
- • **Consistency:** How to have the same view in the face of changes
- • **Fault Tolerance:** How to survive in a world

Characteristics of DS

- Scalability
- Support Heterogeneity
- Continuous Availability
- Users are unaware about the distribution and heterogeneity of the system
- The components of a system are also unaware of the heterogeneity
- Users and applications can interact with the system uniformly

12

Goals of DS

- Connecting Users and Resources
- Transparency
- Openness
- Scalability

13

Connecting Users and Resources

- The main goal is to make easy for users to access remote resources and to share them with other users in a controlled way
 - Security
 - From external users: (Authentication, Firewall)
 - From internal users: (Policies, Log Files)

14

Transparency

- To hide the fact that its processes and resources are physically distributed across multiple computers.

OR

- A distributed system that is able to present itself to users and applications as it were only a single computer system is said to be transparent.

15

Types of Transparency

- Access (Data representation: little endian / big endian)
 - To hide the details of data representations, code translation etc.
- Location (URL)
 - To hide the actual location of the resource.
- Migration
 - To hide the fact that the resource has migrated from previous location.
- Relocation
 - To hide the fact that resource has been reallocated due to non availability from previous location.

16

Types of Transparency

- Replication
 - To hide the fact that resource is copied at different locations to make availability of the resource faster.
- Concurrency
 - This is a type of transparency that is related to concurrent access to same resource.
- Failure
 - To hide any of the failure that occur in accessing the resources.

17

Openness

- An open distributed system support heterogeneity.
- An open distributed system offer services according to standard rules (protocols).
- Interoperability
 - It is an extent by which two implementations of system or components from different manufacturers can co exists.
- Portability
 - An application developed on one system can be executed without modification on another.

18

Scalability

- A distributed system should be expandable
 - Scalable with respect to Size
 - Increase in number of nodes.
 - Scalable Geographically
 - Increase in size with respect to geographical location.
 - Scalable Administratively
 - If we expand the system then administration should not become difficult.

19

Scalability

- Problems to be taken care of in scalability:
 - Centralized Services
 - Services available at central machine i-e server.
 - Centralized Data
 - Data available at central machine i-e data base server.
 - Centralized Algorithms.
 - The algorithms that are running on server machine

20

Parallel Processing or Computing

- Parallel computing
 - Using parallel computer to solve single problems faster
- Parallel computer
 - Multiple-processor system supporting parallel programming
- Parallel programming
 - Programming in a language that supports concurrency explicitly

21

Evolution of Supercomputing

- Supercomputers – Most powerful computers that can currently be built. Note: This definition is time dependent.
- Uses during World War II
 - Hand-computed artillery tables
 - Need to speed computations
 - Army funded ENIAC to speedup calculations
- Uses during the Cold War
 - Nuclear weapon design
 - Intelligence gathering
 - Code-breaking

22

Supercomputer

- General-purpose computer
- Solves individual problems at high speeds, compared with contemporary systems
- Typically costs \$10 million or more
- Traditionally found in government labs

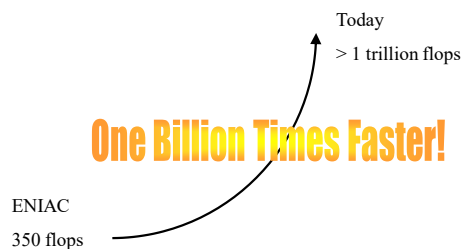
23

Commercial Supercomputing

- Started in capital-intensive industries
 - Petroleum exploration
 - Automobile manufacturing
- Other companies followed suit
 - Pharmaceutical design
 - Consumer products

24

50 Years of Speed Increases



25

CPUs 1 Million Times Faster

- Faster clock speeds
- Greater system concurrency
 - Multiple functional units
 - Concurrent instruction execution
 - Speculative instruction execution

26

Section 2

Hardware Concepts of Distributed Systems

27

Hardware Concepts

- All distributed systems consist of multiple CPUs.
- we can organize hardwares in several different ways in term of their connection and communication
- Various classification schemes have been proposed by the time but none of them have been widely adopted

28

Classification in term of Hardware

- We can classify distributed systems regarding hardware in two broad categories.
 - Multi Processor Systems.
 - Multi Computer Systems.

29

Hardware Requirements

- We cant define complete hardware requirements for a distributed system in general but we will discuss the most basic requirements here and these are.
 - Processors.
 - Memory (Specially RAM).
 - Interconnecting Resources.

30

Multi Processor Systems

- Such systems consist of computer having multiple processors and memory that are connected through a high speed back plane over the mother board.

OR

- A simple configuration is to have a high speed mother board into which CPU and memory cards can be inserted.

31

Multi Processor Systems

- There are two types of multiprocessor system with respect to Memory.
 - Multi Processor System with Shared Memory.
 - Multi Processor System with Non Shared Memory.
- There are two types of multiprocessor system with respect to Inter connection of Memory and Processors.
 - Bus Based Systems.
 - Switch Based Systems.

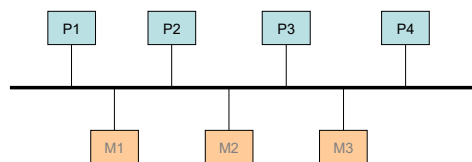
32

Multi Processor Systems With Shared Memory

- In such system there are multiple processors using memory that is shared among the processors.
- Coherent System:
 - In shared memory system since there is only one memory if processor A writes some thing at memory address 40, and after some time processor B reads from memory address 40 it will get the same data written by processor A. Inter process communication is very easy.

33

Multi Processor Systems With Shared Memory



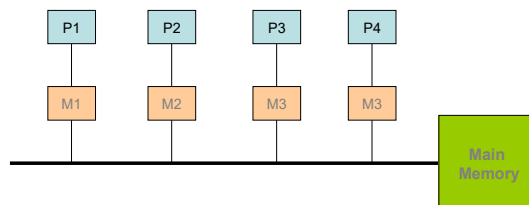
34

Multi Processor Systems With Non Shared Memory

- In such system there are multiple processors each having its own memory along with the main memory.
- Such Systems are incoherent.
- Inter process communication is hard to implement due to incoherence.
- Difficult to build such systems.
- Message Passing Technique is then used to enable inter process communication.
- Generally slow as compare to shared memory systems.
- Advantage is we can connect more Processors.

35

Multi Processor Systems With Non Shared Memory



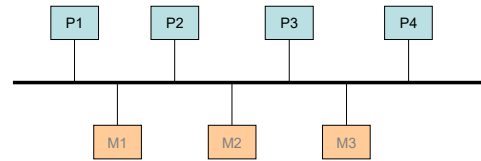
36

Types of Multi Processor Systems With Respect to Interconnection

- There are two types.
 - Bus Based System.
 - Switch Based System.

37

Bus Based Multi Processor Systems With Shared Memory



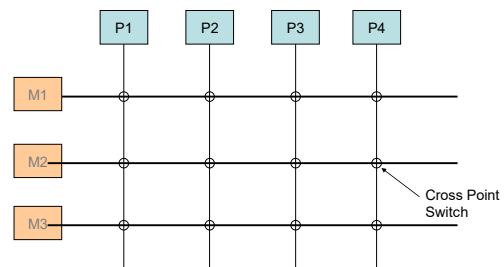
38

Switch Based Multi Processor Systems With Shared Memory

- There are two types.
 - Cross Bar Switch Based.
 - Omega Switch Based.

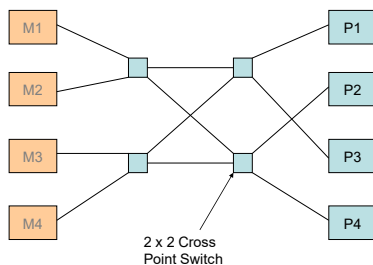
39

Cross Bar Switch Based Multi Processor Systems With Shared Memory



40

Omega Switch Based Multi Processor Systems With Shared Memory



41

Multi Computer Systems

- The system consists of multiple computers connected by interconnecting resources.
- For Example:
 - We have ten computers in the lab, we have ten processors with memories available for process an application.
 - Why not we use them as a single system to process an application.
 - Internet.

42

Multi Computer Systems

- Such system will be incoherent in nature as it will have ten processors, all of which have its own memory.
- In other words No Shared memory.
- Inter Process communication will then achieved through message passing techniques.
- This technique is little bit hard to implement but not as much costly as to buy minicomputer or any other high rank systems.
- And then make program using parallel programming.

43

Types Of Multi Computer Systems

- There are two types based on types of computers we wanted to use in our system
 - Homogeneous Multi Computer systems.
 - Heterogeneous Multi Computer Systems.

44

Types Of Multi Computer Systems

- There are two types further based on Interconnection resources.
 - Homogeneous Multi Computer Systems
 - Bus Based Systems.
 - Switch Based systems.
 - Heterogeneous Multi Computer Systems
 - Bus Based Systems.
 - Switch Based systems.

45

Homogeneous Multi Computer Systems

- The system consist of same type of computers.
- Each Computer has its own memory.
- Computer will communicate each other through a high speed interconnection network.
- This high speed interconnection network can be a bus based or switch based.

46

System Area Network

- SAN are homogeneous multi computer systems based on fast bus based interconnecting network.
- Normally these computers are connected through a shared multi access network (Fast Ethernet).
- They have limited scalability.

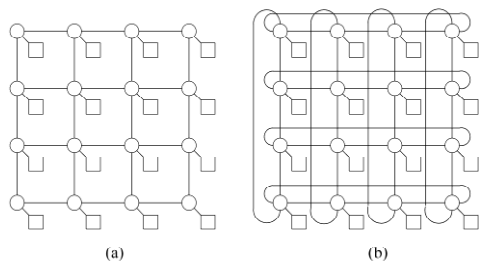
47

Cluster of Workstation

- COWs are homogeneous multi computer systems based on switches.
- Topologies used are
 - Meshes.
 - Hypercubes.

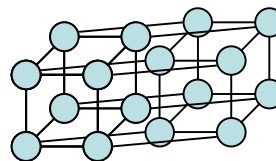
48

Meshes



Note: Circles represent switches and squares represent processors in all these slides.

49



50

Heterogeneous Multi Computer Systems

- The system consist of computers of different types.
- Each Computer has its own memory.
- Computer will communicate each other through a high speed interconnection network.
- This high speed interconnection network can be a bus based or switch based.

51

Bus based Heterogeneous Multi

- These are heterogeneous multi computer systems based on fast bus based interconnecting network.
- Normally these computers are connected through a shared multi access network (Fast Ethernet).
- They have limited scalability.

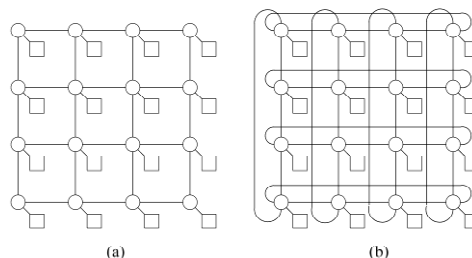
52

Switch Based Heterogeneous Multi Computer Systems

- These are heterogeneous multi computer systems based on switches.
- Topologies used are
 - Meshes.
 - Hypercubes.

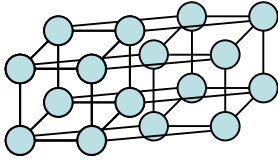
53

Meshes



Note: Circles represent switches and squares represent processors in all these slides.

54



55